

SKRIPSI

PERBANDINGAN ALGORITMA *RANDOM FOREST* DAN *K-NEAREST NEIGHBOR* UNTUK KLASIFIKASI TINGKAT KEPUASAN MAHASISWA TERHADAP PELAYANAN FAKULTAS

COMPARISON OF RANDOM FOREST AND K-NEAREST NEIGHBOR ALGORITHMS FOR CLASSIFICATION OF STUDENT SATISFACTION LEVELS TOWARDS FACULTY SERVICES



MUH. ADRIAN

D0220375

PROGRAM STUDI INFORMATIKA

FAKULTAS TEKNIK

UNIVERSITAS SULAWESI BARAT

MAJENE

2024

HALAMAN PERSETUJUAN

SKRIPSI

**PERBANDINGAN ALGORITMA *RANDOM FOREST* DAN K-NEAREST
NEIGHBOR UNTUK KLASIFIKASI TINGKAT KEPUASAN MAHASISWA
TERHADAP PELAYANAN FAKULTAS**

Telah dipersiapkan dan disusun oleh

Muh. Adrian

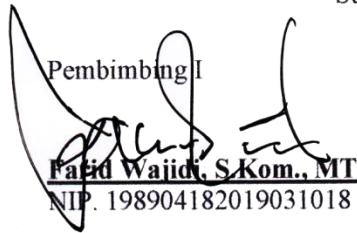
D0220375

Telah dipertahankan di depan Tim Penguji

Pada tanggal, 31 Oktober 2024

Susunan Tim Penguji

Pembimbing I



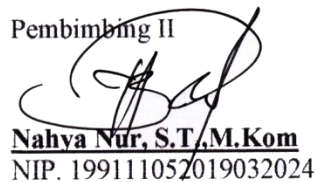
Fahid Wajidi, S.Kom., MT
NIP. 198904182019031018

Penguji I



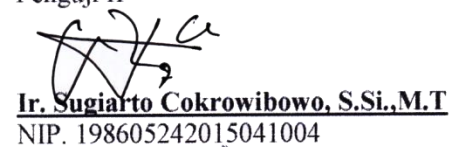
Dr. Eng. Sulfavanti, S.Si., M.T
NIP. 198903172020122011

Pembimbing II



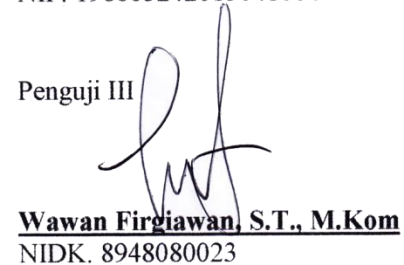
Nahya Nur, S.T., M.Kom
NIP. 199111052019032024

Penguji II



Ir. Sugiarto Cokrowibowo, S.Si., M.T
NIP. 198605242015041004

Penguji III



Wawan Firgiawan, S.T., M.Kom
NIDK. 8948080023

LEMBAR PENGESAHAN SKRIPSI

**PERBANDINGAN ALGORITMA *RANDOM FOREST* DAN K-NEAREST
NEIGHBOR UNTUK KLASIFIKASI TINGKAT KEPUASAN MAHASISWA
TERHADAP PELAYANAN FAKULTAS**

Disusun dan diajukan oleh:

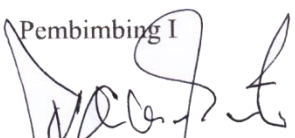
Muh. Adrian

D0220375

Telah dipertahankan di hadapan Panitia Ujian yang dibentuk dalam rangka
Penyelesaian Studi Program Sarjana Teknik Informatika Fakultas Teknik
Universitas Sulawesi Barat
pada tanggal 31 Oktober 2024
dan dinyatakan telah memenuhi syarat kelulusan.

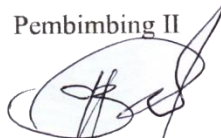
Menyetujui,

Pembimbing I



Farid Wajidi, S.Kom., M.T
NIP. 198904182019031018

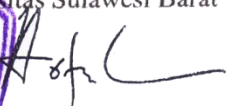
Pembimbing II



Nahya Nur, S.T., M.Kom
NIP. 199111052019032024



Dekan Fakultas Teknik,
Universitas Sulawesi Barat



Dr. Hafsah Nirwana, M.T
NIP. 19640405199003200



Ketua Program Studi
Informatika,



Muh. Rafli Rasyid, S.Kom., M.T
NIP. 198808182022031006

ABSTRAK

Muh. Adrian, Perbandingan Algoritma *Random Forest* Dan *K-Nearest Neighbor* Untuk Klasifikasi Tingkat Kepuasan Mahasiswa Terhadap Pelayanan Fakultas (dibimbing oleh **Farid Wajidi, S.Kom., M.T,** and **Nahya Nur, S.T., M.Kom**)

Penelitian ini membandingkan performa dua algoritma machine learning, yaitu *K-Nearest Neighbors (KNN)* dan *Random Forest (RF)*, dalam klasifikasi tingkat kepuasan mahasiswa terhadap pelayanan fakultas. Klasifikasi dilakukan dengan variasi jumlah tetangga (*K*) pada *KNN* dan jumlah pohon pada *RF*, serta rasio pembagian data latih dan uji. Hasil pengujian menunjukkan bahwa *KNN* menghasilkan akurasi yang sangat tinggi sebesar 99%, dengan presisi 98% untuk kelas puas dan 100% untuk kelas tidak puas. Nilai recall *KNN* juga sangat baik, mencapai 100% untuk kelas puas dan 95% untuk kelas tidak puas. Sebaliknya, *RF* mencatatkan akurasi sebesar 95%, dengan presisi 94% untuk kelas puas dan 95% untuk kelas tidak puas. Namun, nilai recall *RF* menunjukkan variasi, yaitu 90% untuk kelas puas dan 84% untuk kelas tidak puas, yang mengindikasikan bahwa meskipun *RF* memberikan prediksi yang handal, terdapat potensi untuk meningkatkan sensitivitas model, terutama dalam mendeteksi seluruh kelas positif. Perbandingan ini menunjukkan bahwa *KNN* memiliki keunggulan dalam hal akurasi, presisi, dan konsistensi hasil. Sementara itu, *RF* tetap menunjukkan kinerja yang solid, meskipun membutuhkan penyesuaian lebih lanjut untuk meningkatkan sensitivitas pada beberapa kondisi. Hasil penelitian ini memberikan wawasan berharga tentang keandalan kedua algoritma dalam klasifikasi kepuasan mahasiswa, serta potensi perbaikan yang diperlukan, khususnya pada model *Random Forest*.

Kata Kunci: *Machine Learning, Random Forest, K-Nearest Neighbor*, Klasifikasi, Kepuasan Mahasiswa

ABSTRACT

Muh. Adrian, *Comparison Of Random Forest And K-Nearest Neighbor Algorithms For Classification Of Student Satisfaction Levels Towards Faculty Services (supervised by Farid Wajidi, S.Kom., M.T, and Nahya Nur, S.T., M.Kom)*

This study compares the performance of two machine learning algorithms, K-Nearest Neighbors (KNN) and Random Forest (RF), in classifying student satisfaction levels with faculty services. Classification is performed with variations in the number of neighbors (K) for KNN and the number of trees for RF, as well as the training and testing data split ratio. The results show that KNN achieves a very high accuracy of 99%, with precision of 98% for the satisfied class and 100% for the dissatisfied class. KNN's recall is also excellent, reaching 100% for the satisfied class and 95% for the dissatisfied class. In contrast, RF recorded an accuracy of 95%, with precision of 94% for the satisfied class and 95% for the dissatisfied class. However, RF's recall values showed some variation, with 90% for the satisfied class and 84% for the dissatisfied class. This indicates that although RF provides reliable predictions, there is potential for improving the model's sensitivity, especially in detecting all positive class instances. The comparison suggests that KNN has advantages in terms of accuracy, precision, and consistency. Meanwhile, RF still performs solidly, although it requires further adjustments to enhance sensitivity in certain conditions. The findings of this study provide valuable insights into the reliability of both algorithms in classifying student satisfaction, as well as the potential improvements needed, particularly for the Random Forest model.

Keywords: *Machine Learning, Random Forest, K-Nearest Neighbor, Classification, Student Satisfaction*

BAB I

PENDAHULUAN

A. Latar Belakang

Fakultas Teknik Universitas Sulawesi Barat memiliki tiga program studi jenjang strata 1 yaitu Program Studi Teknik informatika, Program Studi Teknik Sipil dan Program Studi Perencanaan Wilayah dan Kota. Sebagai penyelenggara pelayanan pendidikan tinggi yang memprioritaskan pelayanan maksimal Fakultas Teknik Universitas Sulawesi Barat berkomitmen untuk memastikan kepuasan mahasiswa di setiap aspek program studi yang ditawarkan.

Kepuasan mahasiswa adalah salah satu faktor yang sangat penting untuk mendatangkan manfaat bagi fakultas. Seperti halnya industri jasa lainnya kepuasan pelanggan (mahasiswa) memiliki peranan penting dalam memenangkan sebuah persaingan. Menurut Millind M. Lele dan Jagdish N. Sheth (1995) yang dikutip (Mariana et al., 2020) disampaikan bahwa perusahaan-perusahaan yang sukses selalu menempatkan kepuasan pelanggan pada prioritas paling tinggi.

Pengukuran tingkat kepuasan mahasiswa adalah sebuah kegiatan yang seharusnya dilakukan secara berkala untuk mengetahui berapa tingkat pelayanan yang telah diberikan. Jika pelayanan buruk maka pihak universitas wajib untuk dapat meningkatkan pelayanan. Dengan mengetahui tingkat kepuasan mahasiswa, pihak perguruan tinggi dapat memperbaiki layanan apabila kurang baik dan lebih meningkatkan pelayanan apabila tingkat kepuasan sudah baik. Menyikapi anggapan tersebut, maka penulis merasa perlu untuk melakukan analisis kepuasan mahasiswa terhadap pelayanan akademik yang ada (Saragih et al., 2021).

Menurut (Mariana et al., 2020) kepuasan mahasiswa menjadi salah satu faktor penting bagi fakultas dalam rangka memenangkan persaingan dari para pesaing dalam menyelenggarakan program-program yang sama. Yang lebih penting lagi kepuasan mahasiswa dalam keseluruhan pengalaman belajar (*learning*

experience) merupakan komitmen profesional yang harus senantiasa diupayakan *keberwujudannya* sebagai lembaga pendidikan yang bermutu. Fakultas dituntut untuk mampu memberikan kualitas layanan yang dinilai baik dan positif oleh mahasiswa selama mereka menempuh studi baik yang sifatnya *tangible* maupun *Intangible*.

Hasil pengukuran tingkat kepuasan mahasiswa perlu dilakukan klasifikasi. Dengan melakukan klasifikasi, institusi dapat memperoleh gambaran yang jelas mengenai tingkat kepuasan mahasiswa terhadap berbagai aspek layanan akademik dan non-akademik yang diberikan. Hal ini memungkinkan institusi untuk mengidentifikasi area-area yang membutuhkan perbaikan secara spesifik, seperti kualitas pengajaran, fasilitas kampus, atau layanan administrasi. Selain itu, klasifikasi juga membantu institusi untuk memprioritaskan upaya peningkatan kualitas layanan dengan fokus pada aspek-aspek yang dianggap paling penting oleh mahasiswa namun memiliki tingkat kepuasan yang rendah. Dengan adanya klasifikasi, institusi dapat memantau tren kepuasan mahasiswa dari waktu ke waktu serta melakukan benchmarking dengan institusi lain, sehingga memungkinkan evaluasi dan perbaikan yang berkelanjutan dalam rangka meningkatkan kepuasan mahasiswa secara menyeluruh dan memperkuat daya saing institusi.

Penggunaan algoritma atau metode dalam melakukan proses klasifikasi data adalah suatu keharusan untuk mengatasi kompleksitas dan skala data yang semakin meningkat. Algoritma klasifikasi memungkinkan data dikelompokkan secara otomatis berdasarkan karakteristik atau pola yang dimiliki, sehingga dapat meningkatkan efisiensi dan akurasi dalam proses pengambilan keputusan. Kemampuan algoritma untuk beradaptasi dan belajar dari data baru juga menjadikannya semakin relevan dalam era big data saat ini, di mana volume data terus berkembang secara eksponensial.

Pemanfaatan algoritma klasifikasi, seperti *K-Nearest Neighbors* (KNN) dan *Random Forest* (RF), sangat penting dalam menangani kompleksitas dan skala data

yang meningkat. Algoritma ini membantu dalam mengungkap pola tersembunyi dan hubungan signifikan dalam data, meningkatkan akurasi prediksi dan konsistensi hasil analisis (Teoh & Rong, 2022). Dengan menggunakan algoritma klasifikasi, peneliti dan praktisi dapat memperoleh keputusan yang lebih tepat dari wawasan yang dihasilkan data (Esmaeel, 2020). Selanjutnya, algoritma ini merampingkan proses klasifikasi, mempercepat analisis data, dan meningkatkan generalisasi ke kumpulan data baru. Secara keseluruhan, integrasi algoritma dalam klasifikasi data sangat penting untuk memajukan pemahaman berbasis data, prediksi, dan proses pengambilan keputusan.

Penelitian ini diangkat berdasarkan penelitian sebelumnya (Louis Madaerdo Sotarjua & Dian Budhi Santoso, 2022) Berdasarkan hasil performa model klasifikasi dari model algoritma yang digunakan pada penelitian ini, maka model KNN memiliki hasil performa yang terbaik nilai metrik evaluasinya. Sehingga pada penelitian ini, model KNN adalah model klasifikasi yang lebih baik digunakan pada penelitian ini dengan Akurasi 86.57 %, dibandingkan dengan algoritma *Decision Tree* dengan nilai akurasinya 85.29 % dan *Random Forest* dengan akurasinya 86.37 %. Namun pada penelitian (Wibisono & Fahrurrozi, 2019) Berdasarkan perbandingan hasil ujicoba, performance measure algoritma *Random forest* memiliki hasil yang lebih baik dibanding dengan algoritma *Naïve Bayes*, *K-Nearest Neighbor* dan *Decision Tree* dengan metode *k-fold cross validation*. Algoritma *Random Forest* dapat memberikan rerata hasil akurasi sebesar 85,668 % dengan *recall* kelas '1' adalah 89 %, *recall* kelas '0' adalah 83,6%, presisi kelas '1' adalah 85%, dan presisi kelas '0' adalah 85,8%. Dengan demikian dapat dikatakan bahwa algoritma *Random Forest* adalah algoritma yang terbaik dalam mengklasifikasi kasus penyakit jantung koroner dibanding dengan algoritma *Naïve Bayes*, *K-Nearest Neighbor* dan *Decision Tree*.

Berdasarkan latar belakang diatas penulis tertarik untuk mengangkat judul **“Perbandingan Algoritma *Random Forest* dan *K-Nearest Neighbor* Untuk**

Klasifikasi Tingkat Kepuasan Mahasiswa”. *Dataset* yang di analisis yaitu Data Kuesioner mahasiswa teknik universitas Sulawesi barat, yang merupakan hasil data survei tentang kepuasan mahasiswa fakultas teknik universitas Sulawesi barat, *Dataset* ini digunakan untuk mengetahui membandingkan algoritma *Random Forest* dan *K-Nearest Neighbor*. Penelitian ini diharapkan dapat memberikan pemahaman yang lebih mendalam tentang keunggulan dan kelemahan algoritma *Random Forest* dan *K-Nearest Neighbor* serta yang mana lebih baik dalam klasifikasi tingkat kepuasan mahasiswa fakultas teknik. Sehingga dapat memberikan panduan praktis bagi pihak Fakultas dalam meningkatkan pelayanan untuk Mahasiswa.

B. Rumusan Masalah

Berdasarkan latar belakang, adapun rumusan masalah dalam penelitian ini yaitu: Bagaimana hasil Perbandingan antara pendekatan algoritma *Random Forest* dan *K-Nearest Neighbor* dalam klasifikasi tingkat kepuasan Mahasiswa?

C. Tujuan Penelitian

Untuk mengetahui hasil Perbandingan antara pendekatan algoritma *Random Forest* dan *K-Nearest Neighbor* dalam klasifikasi tingkat kepuasan Mahasiswa?

D. Manfaat Penelitian

1. Untuk Instansi Fakultas Teknik: penelitian ini diharapkan dapat memberikan kontribusi positif terhadap perkembangan dan peningkatan kualitas layanan secara keseluruhan. pemahaman yang lebih baik tentang faktor-faktor yang mempengaruhi kepuasan Mahasiswa Teknik dapat menjadi keunggulan kompetitif yang signifikan bagi Pihak Fakultas Teknik.
2. Untuk Pendidikan: memberikan informasi kepada bidang penelitian *forecasting* tentang algoritma apa yang paling baik dari kedua algoritma di

atas dalam kasus mengklasifikasi tingkat kepuasan Mahasiswa pada pelayanan Fakultas Teknik. Selain itu, penelitian ini bisa menjadi dasar bagi penelitian lanjutan di bidang kecerdasan buatan, statistik, dan analisis data.

E. Batasan Masalah Penelitian

Batasan masalah dalam penelitian ini adalah sebagai berikut:

1. Data yang digunakan dalam penelitian ini merupakan data dari Kuesioner yang berupa penilaian kepuasan Mahasiswa Fakultas Teknik Universitas Sulawesi Barat.
2. Algoritma yang digunakan dalam penelitian ini adalah *Random Forest* dan *K-Nearest Neighbor* untuk mengklasifikasi tingkat kepuasan Mahasiswa Fakultas Teknik Universitas Sulawesi Barat.
3. Dalam mengevaluasi model kedua algoritma di atas menggunakan *Confusion Matrix*.

BAB II

TINJAUAN PUSTAKA

A. Landasan Teori

1. Kepuasan Mahasiswa

Mahasiswa adalah sekelompok intelektual muda dalam masyarakat, yang tentunya mengetahui berbagai fasilitas masyarakat yang disediakan oleh pemerintah, mahasiswa memiliki tanggung jawab moral terhadap fasilitas tersebut karena masyarakat didanai oleh uang rakyat. Karena itu Salah satu kewajiban mahasiswa adalah mengabdikan diri pada waktu kuliah mereka untuk mencari perbaikan di masyarakat sekitar. Misalnya, mengikuti kegiatan yang melibatkan kerjasama antara Kementerian Pendidikan Kebudayaan Riset dan Teknologi dengan Perguruan Tinggi yakni dalam program Kampus Merdeka. (Windarti & Suradi, 2019). Menurut (Mariana et al., 2020), Kepuasan mahasiswa adalah salah satu faktor yang sangat penting untuk mendatangkan manfaat bagi fakultas. Seperti halnya industri jasa lainnya kepuasan pelanggan (mahasiswa) memiliki peranan penting dalam memenangkan sebuah persaingan.

2. Tingkat Kepuasan Mahasiswa

Menurut (Saragih et al., 2021), Pengukuran tingkat kepuasan mahasiswa adalah sebuah kegiatan yang seharusnya dilakukan secara berkala untuk mengetahui berapa tingkat pelayanan yang telah diberikan. Jika pelayanan buruk maka pihak universitas wajib untuk dapat meningkatkan pelayanan. Dengan mengetahui tingkat kepuasan mahasiswa, pihak perguruan tinggi dapat memperbaiki layanan apabila kurang baik dan lebih meningkatkan pelayanan.

Kepuasan adalah tingkat emosi yang dirasakan seseorang setelah

membandingkan Kinerja suatu produk atau layanan dengan harapannya. Ketika persepsi sesuai dengan keinginan konsumen, timbul kesenangan, kepuasan, dan loyalitas terhadap yang bersangkutan dan sebaliknya indeks kepuasan, Tiga skenario yang mungkin muncul pada penumpang, yaitu:

1. Pelayanan yang diberikan sesuai dengan harapan pelanggan.
2. Kualitas layanan dan fasilitas yang dapat diakses sesuai dengan kebutuhan pelanggan.
3. Penyajian asli Kinerja organisasi pada berbagai karakteristik utama konsisten dengan asumsi pelanggan.

3. Data Mining

Data mining adalah proses dalam menganalisa data dari perspektif yang berbeda dan menyimpulkannya sebuah informasi-informasi penting yang dapat dipakai untuk keuntungan, memperkecil biaya pengeluaran, atau bahkan keduanya. Secara teknis, data mining dapat disebut sebagai proses untuk menemukan korelasi atau pola dari ratusan atau ribuan field dari sebuah relational database yang besar. Kemampuan Data mining untuk mencari informasi bisnis yang berharga dari basis data yang sangat besar (Saragih et al., 2021). Tujuan utama dari proses Data Mining adalah k mengekstrak atau mengambil inti sari dari pengetahuan dan sekumpulan data sehingga didapatkan struktur yang dapat lebih mudah dimengerti. Dalam penelitian ini, metode data mining yang digunakan adalah klasifikasi (Shafarindu et al., 2021). Menurut (Sumantri & Utami, 2020) beberapa proses yang terdapat dalam penambangan data adalah sebagai berikut:

1. Data Selection

Pemilihan (seleksi) data dari sekumpulan data operasional perlu dilakukan sebelum tahap penggalan informasi dalam *knowledge data discovery* (KDD) dimulai. Data hasil seleksi yang akan digunakan untuk

proses data mining, disimpan dalam suatu berkas, terpisah dari basis data operasional.

2. *Preprocessing/ Cleaning*

Tahap preprocessing/cleaning merupakan proses yang dilakukan dengan cara memeriksa kembali data siswa antara lain untuk pembersihan data-data redudansi, outlier, dan bernilai null (data yang kosong). Hal ini dilakukan agar *input* data yang akan diolah merupakan data yang telah “bersih” sehingga hasil dari perhitungan algoritma data mining pun akan memberikan hasil yang sesuai (Widaningsih, 2022).

3. *Data Transformation*

Transformasi merupakan sekumpulan instruksi untuk merubah *input* menjadi *output* yang diinginkan (*input-proses-output*). Proses yang dilakukan adalah melakukan transformasi data menjadi bentuk yang diinginkan, disesuaikan dengan algoritma klasifikasi yang akan digunakan. Untuk data yang bersifat numerik dilakukan normalisasi untuk menstandarisasi perbedaan skala yang dapat memberikan efek pada hasil yang diperoleh. Data yang bersifat numerik yaitu variabel Nilai Produktif X, Nilai Produktif XI, dan presensi. Untuk itu dilakukan normalisasi data dengan metode minmax (Widaningsih, 2022).

4. *Data Mining*

Data mining adalah proses mencari pola atau informasi menarik dalam data terpilih dengan menggunakan teknik atau metode tertentu. Teknik, metode, atau algoritma dalam data mining sangat bervariasi. Pemilihan metode atau algoritma yang tepat sangat tergantung pada tujuan dan proses KDD secara keseluruhan (Sumantri & Utami, 2020).

5. *Interpretation/ Evaluation*

Pada tahap ini dilakukan evaluasi terhadap hasil klasifikasi.

4. Machine Learning

Machine Learning (ML) adalah disiplin ilmu yang menugaskan komputer untuk belajar dan bertindak seperti manusia, dan meningkatkan kemampuan belajarnya dari waktu ke waktu secara otomatis, dengan cara menyuplai data dan informasi sebagai bentuk pengalaman dan interaksi dunia nyata. Itulah kenapa, salah satu ciri khas dari ML adalah adanya proses pelatihan, pembelajaran, atau training. Oleh karena itu, ML membutuhkan data untuk dipelajari yang disebut sebagai data training (Yudana et al., 2023).

5. Klasifikasi

Klasifikasi merupakan salah satu tahap penting dalam data mining. Klasifikasi adalah pengelompokan data atau objek baru ke dalam kelas atau label berdasarkan atribut-atribut tertentu. Teknik dari klasifikasi adalah dengan melihat variabel dari kelompok data yang sudah ada. Klasifikasi bertujuan untuk memprediksi kelas dari suatu objek yang tidak diketahui sebelumnya. Klasifikasi terdiri dari tiga tahap, yaitu pembangunan model, penerapan model, dan evaluasi. Pembangunan model adalah membangun model menggunakan data latih yang telah memiliki atribut dan kelas. Kemudian, data-data tersebut diterapkan untuk menentukan kelas dari data atau objek yang baru. Setelah itu, data dievaluasi untuk melihat tingkat akurasi dari pembangunan dan penerapan model terhadap data baru (Nasution et al., 2019).

6. Random Forest

Random forest adalah kumpulan dari pohon klasifikasi hasil dari sampling bootstrap data. Langkah awal dalam membangun model *random forest* yaitu menentukan nilai N sebagai jumlah *decision tree* yang dibangun. Pemilihan data yang digunakan untuk pembangunan *tree* menggunakan teknik *bootstrap sample*. Teknik ini akan memilih sampel dari data secara acak dan

dilakukan secara berulang hingga jumlah sampel pada *bootstrap sample* sama dengan jumlah data sebenarnya. Tujuan pemilihan ini bukan untuk mencari kemungkinan terbaik dari *split* semua data, namun hanya mencari kemungkinan terbaik dari *split* dari gen-gen yang terpilih (Riskha Chairunisa et al., 2020).

Menurut (Sandag, 2020) Pohon keputusan dimulai dengan cara menghitung nilai *entropy* sebagai penentu tingkat ketidakmurnian atribut dan nilai *information gain* untuk menghitung nilai *entropy* digunakan rumus seperti pada persamaan 2.1, sedangkan nilai *information gain* menggunakan persamaan 2.2.

Berikut ini merupakan rumus dari *Random forest* :

$$Entropy(S) = \sum_{i=1}^n -pi \log_2 pi \quad (2.1)$$

Keterangan:

S = Himpunan kasus

n = Jumlah anggota dari S

pi = Proporsi dari Si terhadap S

$$Gain(S, A) = Entropy(S) - \sum_{i=1}^n \frac{|Si|}{|S|} * Entropy(Si) \quad (2.2)$$

Keterangan:

S = Himpunan kasus

A = Atribut

n = Jumlah partisi Atribut A

$|Si|$ = Jumlah kasus pada partisi ke- i

$|S|$ = Jumlah kasus dalam S

Menurut Purwa, Taly (2019) Tahapan *Random Forest* adalah sebagai berikut:

1. Menentukan jumlah *tree* (k) yang akan dibentuk.
2. Pengambilan sampel acak sebanyak observasi (*with replacement*)

pada dataset yang berukuran untuk setiap *tree*.

3. Pada setiap *tree*, dilakukan juga pengambilan subset prediktor sebanyak secara acak. Dimana $m < p$, dengan adalah jumlah variabel *predictor*.
4. Ulangi proses ke-2 dan ke-3 sampai sampai sebanyak *tree*.
5. Pada kasus prediksi, hasil prediksi *Random forest* merupakan nilai rata-rata prediksi dari sebanyak *tree*. Sedangkan pada kasus klasifikasi, hasil prediksi *Random forest* diperoleh dari *vote* terbanyak (*majority vote*) dari hasil klasifikasi sebanyak *tree*.

7. *K-Nearest Neighbor*

Algoritma *K-Nearest Neighbor* merupakan teknik klasifikasi data berdasarkan pada klasifikasi data pada data latih yang memiliki kedekatan dengan data uji, dengan sejumlah k data yang terdekat. Teknik ini menghitung kedekatan kasus baru dengan kasus lama berdasarkan pencocokan bobot. Terdapat tiga kunci dalam algoritma ini yaitu sejumlah variabel, jarak atau kesamaan antar objek, dan ukuran dari k yaitu sejumlah data terdekat. Untuk mengklasifikasikan data yang belum diketahui kelasnya, jarak data baru ke data yang telah diketahui kelasnya dihitung, k -terdekatnya diidentifikasi, dan label kelas terdekat ini kemudian digunakan untuk menentukan label kelas dari objek tersebut (Widaningsih, 2022).

Adapun Rumus pencarian jarak dengan menggunakan rumus euclidean:

$$d_{euclidean(x,y)} = \sqrt{\sum_{i=1}^n (x_i - y_i)^2} \quad (2.3)$$

Dimana:

x_i = data testing ke- i

d = jarak

y_i = data training ke- i

n = dimensi data

i = variable data

8. Evaluasi Model

Evaluasi model adalah sebuah proses untuk mengukur performa model yang telah dihasilkan pada proses training. Ada banyak metode untuk mengevaluasi model, biasanya tergantung kepada jenis *task* yang dilakukan serta informasi apa yang ingin diketahui. Sebagai contoh pada kasus regresi dan klasifikasi, membutuhkan metode yang berbeda karena ukuran baik atau tidaknya model regresi dan klasifikasi berbeda. Metode yang dapat digunakan dalam mengevaluasi model yaitu *Confusion Matrix*.

Dalam penelitian (Surahmat & Sutrisno, 2023), *Confusion matrix* adalah sebuah model pengujian tabel yang memiliki 4 (empat) kombinasi nilai prediksi (*predicted values*) dan nilai aktual (*actual values*) yang terdiri dari TP (*True Positive*) merupakan jumlah data positif yang terklasifikasi dengan benar oleh system, TN (*True Negative*) merupakan data negatif yang terklasifikasi dengan benar oleh sistem, FN (*False Negative*) merupakan data negatif namun terklasifikasi salah oleh sistem, FP (*False Positive*) merupakan data positif namun terklasifikasi salah oleh sistem.

Dalam *confusion matrix* nilai akurasi merupakan perbandingan antara data yang terklasifikasi benar dengan keseluruhan data. Adapun nilai akurasi dihitung dengan persamaan :

$$Akurasi = \frac{TP+TN}{TP+TN+FP+FN} \times 100\% \quad (2.4)$$

Dilanjutkan dengan pencarian nilai presisi yang merupakan jumlah data kategori positif yang diklasifikasi secara benar dibagi dengan total data yang diklasifikasi positif dengan menggunakan persamaan :

$$Presisi = \frac{TP}{TP+FP} \times 100\% \quad (2.5)$$

Untuk nilai *recall* merupakan persen data kategori positif yang terklasifikasi dengan benar oleh sistem dengan menghitung persamaan :

$$Recall = \frac{TP}{TP+FN} \times 100\% \quad (2.6)$$

Dilanjutkan dengan nilai error merupakan kasus yang diidentifikasi salah dalam sejumlah data, sehingga dapat dilihat berapa besaran tingkat kesalahan pada sistem yang sedang digunakan. Untuk menghitung persentase error dapat dilakukan melalui persamaan rumus berikut:

$$Error = \frac{FP}{TP} \times 100\% \quad (2.7)$$

B. Penelitian Terkait

Dalam persiapan penelitian ini, penulis merasa penting untuk merujuk pada beberapa penelitian sebelumnya sebagai landasan dan wawasan. Hal ini bertujuan untuk memperoleh informasi dan ide-ide yang dapat dikembangkan dalam penelitian ini. Untuk itu, penulis telah menghimpun sejumlah referensi dari penelitian-penelitian terdahulu yang relevan dengan permasalahan yang dibahas dalam penelitian ini.

Tabel 2. 1 Penelitian Terkait

No	Nama/ Tahun	Judul	Hasil	Perbedaan	Persamaan
1	Saragih, D A R Safii, M, 2021	Penerapan Data Mining Klasifikasi Tingkat Kepuasan Mahasiswa Terhadap Pelayanan Sistem Informasi di Program Studi Sistem Informasi	Hasil yang didapat dengan adanya analisis kepuasan mahasiswa terhadap pelayanan sistem informasi di STIKOM Tunas Bangsa menggunakan metode C4.5, maka tingkat kepuasan mahasiswa STIKOM Tunas Bangsa dapat terukur dengan jelas. Setelah dilakukan analisis, dari beberapa aspek yang menjadi tolak ukur penilaian kepuasan	Yang menjadi pembeda dalam penelitian ini yaitu penelitian ini menggunakan metode c4.5 untuk klasifikasi tingkat kepuasan mahasiswa sedangkan penelitian yang akan dibuat yaitu perbandingan algoritma KNN dan <i>Random forest</i> dengan data set kepuasan mahasiswa	Yang menjadi Persamaan sama - sama kasus klasifikasi tingkat kepuasan mahasiswa

			<p>mahasiswa maka dapat diketahui bawah aspek kepuasan yang paling dominan dari beberapa aspek kualitas sistem informasi yang diberikan oleh pihak kampus STIKOM Tunas Bangsa</p>		
2	<p>Natuzzuh riyah, Ami Nafisah, Nisa, 2021</p>	<p>Klasifikasi Tingkat Kepuasan Mahasiswa Terhadap Pembelajaran Secara Daring Menggunakan Algoritma <i>Naïve Bayes</i></p>	<p>perhitungan data mining dengan metode klasifikasi menggunakan algoritma <i>naïve bayes</i> kesimpulan yang didapatkan dari data set yang diperoleh dengan 2 <i>class</i> puas dan tidak puas bahwa hasil pengujian algoritma <i>naïve bayes</i> ini diperoleh tingkat akurasi sebesar 76,92% dan class <i>precision</i> menghasilkan 100.00%, class <i>recall</i> 57.14% dan nilai AUC 0.881 mendekati angka 1 jadi model yang dihasilkan baik. Jadi, dari hasil</p>	<p>Yang menjadi pembeda dalam penelitian ini yaitu penelitian ini menggunakan algoritma <i>Naïve Bayes</i> untuk klasifikasi tingkat kepuasan mahasiswa terhadap pembelajaran secara daring sedangkan penelitian yang akan dibuat yaitu perbandingan algoritma KNN dan <i>Random forest</i> dengan data set kepuasan mahasiswa</p>	<p>Yang menjadi Persamaan sama - sama kasus klasifikasi tingkat kepuasan mahasiswa</p>

3	Rofik, M A Siregar, A M, 2021	Perbandingan Tingkat Kepuasan Siswa Terhadap Pelayanan Sekolah Menggunakan Algoritma <i>K-Means</i> Dan <i>K-Medoids</i>	<p>Hasil perhitungan menggunakan algoritma <i>k-means</i> dan <i>k-medoids</i> dengan Kumpulan data dengan 509 data kuesioner yang diisi siswa, dan lima atribut yang digunakan, hasil yang diperoleh <i>k-means</i> setelah 3 iterasi siswa merasa puas dengan pelayanan sekolah berjumlah 276 siswa, cukup puas berjumlah 216 siswa dan kurang puas dengan pelayanan sekolah berjumlah 17 siswa. Sedangkan <i>k-medoids</i> lebih baik karena hanya 2 iterasi dengan hasil siswa merasa puas dengan pelayanan sekolah berjumlah 324 siswa, cukup puas berjumlah 11 siswa dan</p>	<p>Yang menjadi pembeda dalam penelitian ini yaitu perbedaan algoritma yang digunakan dalam penelitian ini membandingkan <i>K-means</i> dan <i>K-Medoids</i> sedangkan penelitian yang akan dibuat yaitu perbandingan algoritma KNN dan <i>Random forest</i> dengan data set kepuasan mahasiswa</p>	<p>Yang menjadi Persamaan sama-sama perbandingan algoritma dengan kasus topic menyerupai</p>
---	-------------------------------	--	--	---	--

			<p>kurang puas dengan pelayanan sekolah berjumlah 174 siswa, sehingga algoritma k-medoids lebih cocok digunakan untuk klustering kepuasan siswa disekolah.</p>		
4	<p>Widiastuti, Tri Karsa, Koko Juliane, Christina, 2022</p>	<p>Evaluasi Tingkat Kepuasan Mahasiswa Terhadap Pelayanan Akademik Menggunakan Metode Klasifikasi Algoritma C4.5</p>	<p>hasil klasifikasi untuk jenis variabel yang digunakan untuk mengukur tingkat ketidakpuasan mahasiswa terhadap sistem pendukung akademik yang digunakan sebagai titik awal untuk menentukan apa yang perlu dilakukan un Hasil implementasi hanya menggunakan satu aplikasi di U.J. Algoritma penambangan data C4.5, yang dikenal sebagai RapidMiner Studio, dipilih sebagai hasil evaluasi mahasiswa</p>	<p>Yang menjadi pembeda dalam penelitian ini yaitu penelitian ini menggunakan metode c4.5 untuk klasifikasi tingkat kepuasan mahasiswa sedangkan penelitian yang akan dibuat yaitu perbandingan algoritma KNN dan <i>Random forest</i> dengan data set kepuasan mahasiswa</p>	<p>Yang menjadi Persamaan sama - sama kasus klasifikasi tingkat kepuasan mahasiswa</p>

			<p>master terhadap layanan akademik dalam kategori Tangible, Empathy, Responsiveness, Reliability, dan Assurance. Evaluasi ini menghasilkan skor 90%, yang merupakan skor yang sangat tinggi. Studi ini dilakukan untuk memastikan dilakukan secara rutin dan dapat berhasil dengan menambahkan beberapa fitur yang sangat penting untuk menurunkan zona batas kandidat magister. Untuk meningkatkan level bantuan yang diberikan kepada mahasiswa level magister</p>		
5	Fatmawati, Narti, 2022	Perbandingan Algoritma C4.5 dan <i>Naive Bayes</i> Dalam Klasifikasi Tingkat Kepuasan	Hasil penelitian yang diperoleh dari perbandingan algoritma C4.5 dan <i>naive Bayes</i> terhadap	Yang menjadi pembeda dalam penelitian ini yaitu penelitian ini menggunakan perbandingan algoritma C4.5	Yang menjadi Persamaan sama-sama perbandingan algoritma dengan kasus tingkat kepuasan

		Mahasiswa Terhadap Pembelajaran Daring	pembelajaran daring bahwa hasil pengujian dari kedua algoritma tersebut dengan menggunakan <i>software RapidMiner</i> dapat disimpulkan bahwa kinerja dari algoritma Naïve Bayes lebih baik dibandingkan dengan C4.5, hal ini dapat dilihat dari nilai akurasi Algoritma Naïve Bayes dengan nilai 70.59% sedangkan algoritma C4.5 adalah 58.82%	dan Naïve bayes dengan data set kepuasan mahasiswa sedangkan penelitian yang akan dibuat yaitu perbandingan algoritma KNN dan <i>Random forest</i> dengan data set kepuasan mahasiswa	mahasiswa
6	Louis Madaerd o Sotarjua Dian Budhi Santoso, 2022	Perbandingan Algoritma Knn, <i>Decision Tree</i> , Dan <i>Random forest</i> Pada Data <i>Imbalanced Class</i> Untuk Klasifikasi Promosi Karyawan	Hasil performa model klasifikasi dari model algoritma yang digunakan pada penelitian ini, maka model KNN memiliki hasil performa yang terbaik nilai matriks evaluasinya. Sehingga pada penelitian ini, model KNN	Yang menjadi pembeda dalam penelitian ini yaitu topic penelitian sedangkan penelitian yang akan dibuat yaitu perbandingan algoritma KNN dan <i>Random forest</i> dengan data set kepuasan	Yang menjadi Persamaan sama - sama perbandingan algoritma <i>random forest</i> dan knn dengan kasus tingkat kepuasan mahasiswa

			adalah model klasifikasi yang lebih baik digunakan pada penelitian ini, dibandingkan dengan algoritma Decision Tree dan <i>Random forest</i> . Pada penelitian ini juga hasil performa model klasifikasi yang dilakukan tidak ditemukan <i>overfitting</i> maupun <i>underfitting</i> , sehingga model dapat berperforma baik pada <i>training</i> maupun <i>testing</i>	mahasiswa	
7	Raudya, Talitha Irawan, Bambang Faqih, Ahmad, 2023	Perbandingan implementasi 2 algoritma svm dan k-nn Dalam pengklasifikasi an kepuasan pengguna <i>smart e-learning</i>	Dari implementasi dua algoritma menghasilkan kesimpulan bahwa penggunaan algoritma <i>K-Nearest Neighbor</i> lebih baik dikarenakan tingkat akurasi yang lebih tinggi dibandingkan dengan algoritma <i>Support Vector Machine</i> . Algoritma <i>K-</i>	Yang menjadi pembeda dalam penelitian ini yaitu penelitian ini menggunakan perbandingan algoritma KNN dan SVM dengan data set kepuasan pengguna smart e-learning sedangkan penelitian yang akan dibuat yaitu perbandingan	Yang menjadi Persamaan sama - sama perbandingan algoritma KNN dengan kasus tingkat kepuasan.

			<p><i>Nearest Neighbor</i> menghasilkan akurasi sebesar 94,00% dan <i>Support Vector Machine</i> sebesar 56,00%. Uji T-Test menunjukkan nilai signifikansi sebesar 0,000 pada algoritma <i>K-Nearest Neighbor</i> sehingga dapat dinyatakan bahwa algoritma ini merupakan metode yang cocok diterapkan dalam penelitian ini.</p>	<p>algoritma KNN dan <i>Random forest</i> dengan data set kepuasan mahasiswa.</p>	
8	<p>Rahmat, A.W Ladjamuddin, M.S Awaludin, T.D, 2023</p>	<p>Perbandingan Algoritma <i>Decision Tree</i>, <i>Random forest</i> Dan <i>Naive Bayes</i> Pada Prediksi Penilaian Kepuasan Penumpang Maskapai Pesawat Menggunakan Dataset Kaggle</p>	<p>hasil perbandingan performa ketiga algoritma bahwa algoritma <i>random forest</i> memiliki kelebihan nilai akurasi yang didapat pada algoritma ini merupakan yang terbesar diantara dua algoritma pembandingnya. Algoritma <i>decision tree</i> memiliki kelebihan nilai akurasi yang</p>	<p>Yang menjadi pembeda dalam penelitian ini yaitu penelitian ini menggunakan perbandingan 3 algoritma <i>Decision Tree</i>, <i>Random forest</i>, dan <i>Naive Bayes</i> untuk prediksi penilaian kepuasan penumpang maskapai pesawat dengan data set dari situs kaggle</p>	<p>Yang menjadi Persamaan sama - sama perbandingan algoritma <i>Random forest</i> dengan kasus tingkat kepuasan.</p>

			<p>dihasilkan lebih tinggi dari algoritma <i>naïve bayes</i>, kekurangan dari algoritma <i>decision tree</i> adalah nilai akurasi yang dihasilkan masih lebih rendah dari algoritma <i>random forest</i>. Algoritma <i>naïve bayes</i> memiliki kelebihan nilai akurasi yang dihasilkan masih dapat lebih baik seiring dengan pemodelan yang dilakukan, kekurangan dari algoritma ini nilai akurasi yang dihasilkan jauh lebih rendah dari algoritma <i>random forest</i> dan <i>decision tree</i>.</p>	<p>sedangkan penelitian yang akan dibuat yaitu perbandingan algoritma KNN dan <i>Random forest</i> dengan data set kepuasan mahasiswa dari data data kuesioner.</p>	
9	Sulaiman ,Edi Surya Negara, 2023	Komparasi Algoritma <i>K-Nearest Neighbors</i> dan <i>Random forest</i> Pada Prediksi Harga Mobil Bekas	<p>Berdasarkan beberapa skenario pengujian dan analisis yang dilakukan, dapat diambil kesimpulan sebagai berikut, perbandingan antara Algoritma</p>	<p>Yang menjadi pembeda dalam penelitian ini yaitu topik penelitian yang digunakan dalam penelitian ini menggunakan topik Komparasi Algoritma <i>K-</i></p>	<p>Yang menjadi Persamaan sama - sama perbandingan algoritma knn dan <i>random forest</i> dengan kasus topic yang berbeda.</p>

			<p><i>K-Nearest Neighbors</i> dan <i>Random forest</i> mendapatkan hasil nilai error yang cukup jauh. Hal ini dibuktikan dengan proses pengujian dimana KNN mendapatkan nilai 0.009817, <i>Random forest</i> 0.001003. Algoritma yang di sarankan penulis dalam pembuatan model machine learning pada studi kasus prediksi harga mobil bekas ini adalah <i>Random forest</i>. Serta akurasi pada model menunjukkan keunggulan model <i>Random forest</i> yang mendapatkan nilai 96.38 % sedangkan model KNN mendapatkan nilai 59.17 %. Dapat ditarik kesimpulan bahwa model menggunakan</p>	<p><i>Nearest Neighbors</i> dan <i>Random forest</i> Pada Prediksi Harga Mobil Bekas sedangkan penelitian yang akan dibuat yaitu perbandingan algoritma KNN dan <i>Random forest</i> dengan data set kepuasan mahasiswa</p>	
--	--	--	---	---	--

			<p>algoritma <i>Random forest</i> lebih unggul dibandingkan <i>K-Nearest Neighbors</i>.</p>		
10	<p>Fredilio Julfikar Rahma Stiven Hamonangan Sinurat, Daniel Ryan Hamonangan Sitompu, Dennis Jusuf Ziege, Evta Indra, 2023</p>	<p>Perbandingan Algoritma <i>K-Nearest Neighbors</i> (K-NN) dan <i>Random forest</i> terhadap Penyakit Gagal Jantung</p>	<p>Model <i>machine learning</i> yang telah dibuat dapat menentukan faktor apa yang menyebabkan penyakit gagal jantung pada pasien. Berdasarkan hasil dari analisis dua algoritma <i>K-Nearest Neighbor</i> (KNN) dan algoritma <i>Random forest</i>, maka diperoleh hasil akurasi terbaik dari algoritma <i>Random forest</i> sebesar 96,5 %</p>	<p>Yang menjadi pembeda dalam penelitian ini yaitu topik penelitian yang digunakan dalam penelitian ini menggunakan topik penyakit gagal jantung Algoritma <i>K-Nearest Neighbors</i> dan <i>Random forest</i> Pada Prediksi Harga Mobil Bekas sedangkan penelitian yang akan dibuat yaitu perbandingan algoritma KNN dan <i>Random forest</i> dengan data set kepuasan mahasiswa</p>	<p>Yang menjadi Persamaan sama - sama perbandingan algoritma knn dan <i>random forest</i> dengan kasus topic yang berbeda.</p>

DAFTAR PUSTAKA

- Esmaeel, H. R. (2020). Analysis of classification learning algorithms. *Indonesian Journal of Electrical Engineering and Computer Science*, 17(2), 1029–1039. <https://doi.org/10.11591/IJEECS.V17.I2.PP1029-1039>
- LOUIS MADAERDO SOTARJUA, & DIAN BUDHI SANTOSO. (2022). Perbandingan Algoritma Knn, Decision Tree,*Dan Random*Forest Pada Data Imbalanced Class Untuk Klasifikasi Promosi Karyawan. *Jurnal INSTEK (Informatika Sains Dan Teknologi)*, 7(2), 192–200. <https://doi.org/10.24252/instek.v7i2.31385>
- Mariana, N., Redjeki, R. S., & Purwatiningtyas. (2020). Analisis kepuasan mahasiswa terhadap kualitas layanan (service quality) fakultas teknologi informasi jurnal ikra-ith informatika vol 4 no 3 november 2020 jurnal ikra-ith informatika vol 4 no 3 november 2020. *Jurnal IKRA-ITH Informatika*, 4(3), 22–30.
- Nasution, D. A., Khotimah, H. H., & Chamidah, N. (2019). Perbandingan Normalisasi Data untuk Klasifikasi Wine Menggunakan Algoritma K-NN. *Computer Engineering, Science and System Journal*, 4(1), 78. <https://doi.org/10.24114/cess.v4i1.11458>
- Raudya, T., Irawan, B., & Faqih, A. (2023). Perbandingan Implementasi 2 Algoritma Svm Dan K-Nn Dalam Pengklasifikasian Kepuasan Pengguna Smart E-Learning. *JATI (Jurnal Mahasiswa Teknik Informatika)*, 7(3), 2057–2062. <https://doi.org/10.36040/jati.v7i3.7067>
- Riska Chairunisa, Adiwijaya, & Widi Astuti. (2020). Perbandingan CART dan Random Forest untuk Deteksi Kanker berbasis Klasifikasi Data Microarray. *Jurnal RESTI (Rekayasa Sistem Dan Teknologi Informasi)*, 4(5), 805–812. <https://doi.org/10.29207/resti.v4i5.2083>
- Sandag, G. A. (2020). Prediksi Rating Aplikasi App Store Menggunakan Algoritma

- Random Forest. *CogITO Smart Journal*, 6(2), 167–178.
<https://doi.org/10.31154/cogito.v6i2.270.167-178>
- Saragih, D. A. R., Safii, M., & ... (2021). Penerapan Data Mining Klasifikasi Tingkat Kepuasan Mahasiswa Terhadap Pelayanan Sistem Informasi di Program Studi Sistem Informasi. *Journal of Information ...*, 2(2), 173–177.
<http://ejournal.seminar-id.com/index.php/josh/article/view/628%0Ahttp://ejournal.seminar-id.com/index.php/josh/article/download/628/425>
- Shafarindu, A. I., Ernawati, L., & Zaidiah, A. (2021). Penerapan Algoritma Naive Bayes untuk Klasifikasi Tingkat Kebugaran Jasmani Berdasarkan Hasil Pengukuran pada Pegawai. *Seminar Nasional Mahasiswa Ilmu Komputer Dan Aplikasinya (SENAMIKA)*, September, 278–287.
- Sumantri, R. B. B., & Utami, E. (2020). Penentuan Status Tahapan Keluarga Sejahtera Kecamatan Sidareja Menggunakan Teknik Data Mining. *Respati*, 15(3), 71. <https://doi.org/10.35842/jtir.v15i3.375>
- Surahmat, A., & Sutrisno, M. (2023). Analisis Kepuasan Pelanggan Dalam Industri Teknologi Menggunakan Algoritma C4.5. *Analisis Kepuasan Pelanggan Dalam Industri Teknologi Menggunakan Algoritma C4.5*, 13(2), 75–79.
- Teoh, T. T., & Rong, Z. (2022). *Classification BT - Artificial Intelligence with Python* (T. T. Teoh & Z. Rong (eds.); pp. 183–211). Springer Singapore.
https://doi.org/10.1007/978-981-16-8615-3_11
- Wibisono, A. B., & Fahrurozi, A. (2019). Perbandingan Algoritma Klasifikasi Dalam Pengklasifikasian Data Penyakit Jantung Koroner. *Jurnal Ilmiah Teknologi Dan Rekayasa*, 24(3), 161–170. <https://doi.org/10.35760/tr.2019.v24i3.2393>
- Widaningsih, S. (2022). Penerapan Data Mining untuk Memprediksi Siswa Berprestasi dengan Menggunakan Algoritma K Nearest Neighbor. *JATISI (Jurnal Teknik Informatika Dan Sistem Informasi)*, 9(3), 2598–2611.

<https://doi.org/10.35957/jatisi.v9i3.859>

Windarti, M., & Suradi, A. (2019). Perbandingan Kinerja 6 Algoritme Klasifikasi Data Mining untuk Prediksi Masa Studi Mahasiswa. *Telematika*, 12(1), 14. <https://doi.org/10.35671/telematika.v12i1.778>

Yudana, F. R., Suyanto, M., & Nasiri, A. (2023). Model Klasifikasi Untuk Menentukan Kesiapan Kerja Mahasiswa Dan Kelulusan Tepat Waktu Dengan Metode Machine Learning. *IJITECH: Indonesian Journal of Information Technology*, 1(1), 1–12. <https://doi.org/10.37680/ijitech.v1i1.xx>